Abstract

One of the main areas of research in knowledge discovery from data (KDD) is pattern extraction. In this thesis, data mining approaches have been used to generate and extract meaningful patterns from ecological datasets. Various case studies have been performed to extract pattern-based knowledge discovery from species biodiversity datasets. The domain researchers can use the proposed algorithmic frameworks to extract useful information from biodiversity datasets. Also, the proposed methods make the most efficient use of resources to find frequent knowledge patterns, such as frequent itemsets, bi-clusters, and association rules.

Biodiversity and ecosystem informatics is an emerging interdisciplinary field that mitigates the research gap among computer scientists, biologists, conservationists, and others related to various aspects of biological diversity. In this thesis, the study has spawned such initiatives and shown the support of algorithmic applications in the context of ecological improvements.

A few case studies performed on Sundarban mangroves, Indian estuarine data of flora and fauna, estuarine fish datasets, as well as, forest cover across different states of India. A novel framework has been proposed to compute the dark diversity to mitigate biodiversity loss. Another novel framework for studying salt marsh and mangrove species co-existence patterns has been proposed. Two novel algorithms for extracting frequent patterns and frequent closed patterns have been proposed. Both use multiple FP-tree data structures and the concept of cellular learning automata. Detailed experimental analysis concerning the leading algorithms on publicly available benchmark datasets shows the better performance of the suggested approaches. A novel concept has been formulated for extracting constant and coherent signed biclusters. Also, a framework along with a novel domain-specific rule filtering metric has been proposed.